2013

The role of bacterial microcompartments in the fermentation of D-arabinose in Clostridium phytofermentans

Megan A. Strough
University of Massachusetts Amherst

THE ROLE OF BACTERIAL MICROCOMPARTMENTS IN THE FERMENTATION OF D-ARABINOSE IN CLOSTRIDIUM PHYTOFERMENTANS

A Thesis Presented

by

MEGAN A STROUGH

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

MASTER OF MICROBIOLOGY

May 2013

Department of Microbiology

THE ROLE OF BACTERIAL MICROCOMPARTMENTS IN THE FERMENTATION OF D-ARABINOSE IN CLOSTRIDIUM PHYTOFERMENTANS

A Thesis Presented

by

MEGAN A STROUGH

Approved as to style and content by:

_____
Jeffrey Blanchard, Chair

_____
Susan Leschine, Member

_____
Danny Schnell, Member

_____
John M. Lopes, Department Head
Department of Microbiology

ACKNOWLEDGEMENTS

I would like to thank my advisor, Jeffrey L. Blanchard for taking the time to not only guide me through this learning experience but also to share his passion for the study of microbiology. In addition I would like to thank Susan B. Leschine for her continuous support and aid in my professional development. Thanks are also due to Danny J. Schnell, as an active member of my committee, his comments and suggestions were helpful in fulfilling the requirements of this project.

I would like to express my deepest appreciation to all of my lab mates for providing a stable and fundamental knowledge base as well as assisting in the development of current methods and laboratory techniques.

A special thank you to my loving family and significant other who provided the encouragement and moral support necessary to see me through this journey.

ABSTRACT

THE ROLE OF BACTERIAL MICROCOMPARTMENTS IN THE FERMENTATION OF D-ARABINOSE IN CLOSTRIDIUM PHYTOFERMENTANS

MAY 2013

MEGAN A STROUGH, B.S. STATE UNIVERSITY OF NEW YORK AT POTSDAM

M.S., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Jeffrey L. Blanchard

Bacterial microcompartments, or BMCs, are 80-200nm, enzyme-encompassing organelles composed of interlocking proteins that form cyclical hexamers with a small central pore. *Clostridium phytofermentans* or Cphy, is a Gram-positive, rod shaped, anaerobic soil microbe that has the ability to not only break down multiple polysaccharides simultaneously but also proceeds to ferment them into biofuels. The genome of Cphy contains 3 BMC loci. During growth on fucose and rhamnose, one of these loci is highly expressed and microcompartments can be viewed using TEM. Under these growth conditions, three products, ethanol, propanol and propionate, which could potentially be highly useful in the biofuel and bioproducts industries, are produced. Gene expression microarrays have revealed that the genes for the fucose/rhamnose-related microcompartment are also highly expressed on D-arabinose. The role of BMCs during growth on D-arabinose has not been reported in any organism. My goal is to determine the role of BMCs in D-arabinose metabolism in *C. phytofermentans*.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Lignocellulosic plants capture solar energy and convert it into chemical energy by means of a process commonly known as photosynthesis. ($6CO_2 + 6H_2O + Light \rightarrow C_6H_{12}O_6 + 6O_2$) This energy is then stored in the form of carbohydrate polymers such as cellulose and hemicellulose and the aromatic polymer lignin. Modern day plant mass is well endowed with a structural robustness that can be attributed to the evolution of ester and ether linkages between the lignin and polysaccharides present in the cell wall. These vigorous linkages aid in the resistance of lignocellulosic plant mass to degradation by microbial organisms. Not only is cellulose the primary constituent of lignocellulosic plant mass, but it is also the most abundant biopolymer produced in terrestrial environments globally [8]. It is composed of a lengthy chain of ß(1-4)-linked glucose molecules. Hemicellulose, the second most abundant component, is a family of polysaccharides comprised of several 5- and 6-carbon sugars such as xylose, glucose, galactose, mannose, fucose, rhamnose and arabinose. The polymerization of these components allows for the synthesis of lignin, a three-dimensional polymer of phenylpropanoid units that acts as the "cellular glue". The collaboration of lignin, cellulose and hemicellulose results in microfibril structures. These microfibrils are clustered in multiples to produce macrofibrils and it is these macrofibrils that are responsible for the structural recalcitrance of the plant cell wall [12]. (Figure 1)

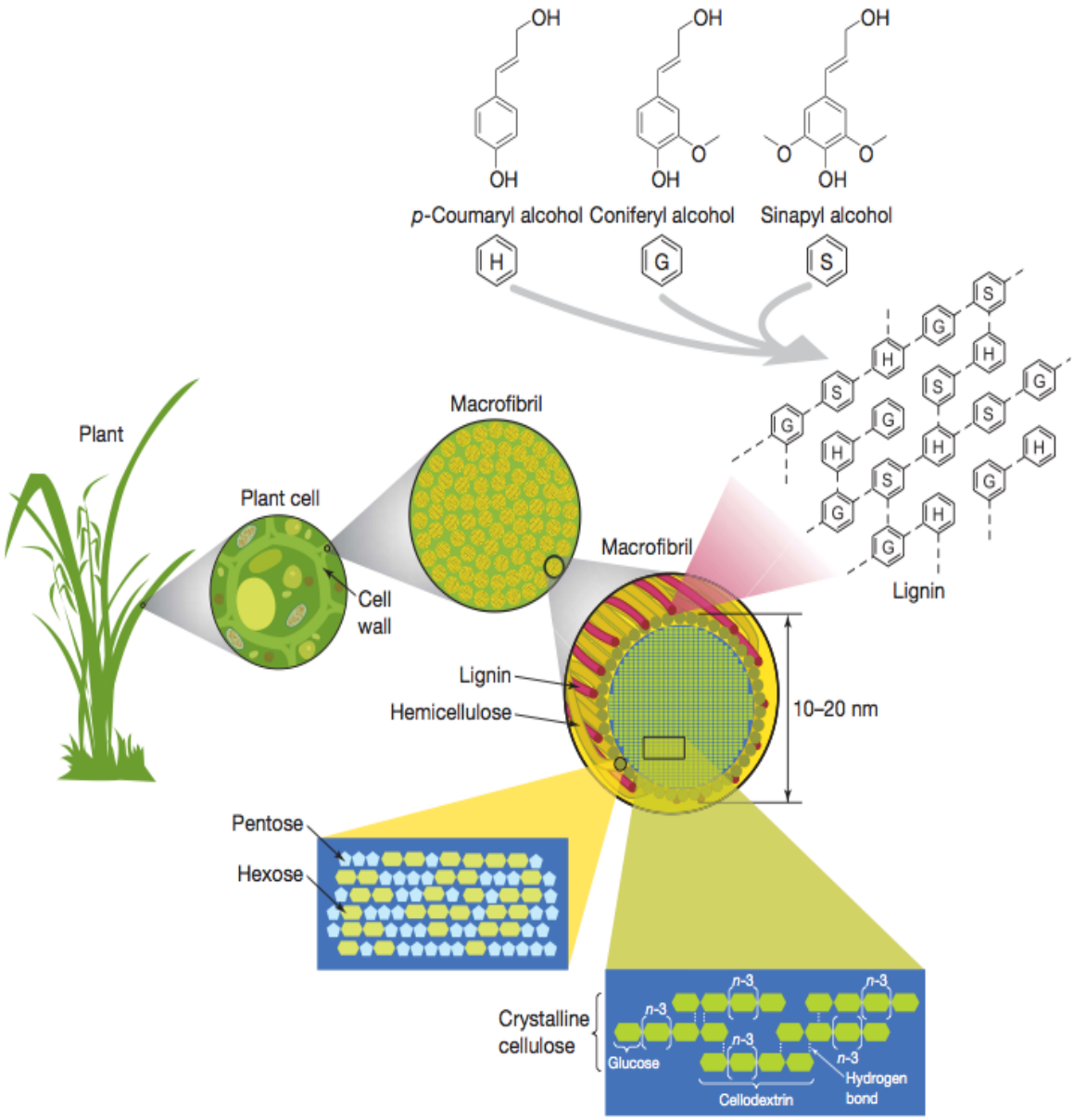

Figure 1. Lignocellulosic Plant Structure. The structural stability of lignocellulosic plant mass is attributed to both the pentose and hexose sugars as well as the ester and ether linkages that fabricate the cell wall [12].

Terrestrial plants conduct photosynthesis by means of three different pathways: C_3, C_4 and crassulacean acid metabolism (CAM). C_3 and C_4 photosynthetic plants are the primary focus of carbon cycle studies due to the fact that CAM plants are not believed to be a significant

contributor to the global carbon cycle. C_3 plants are titled as such because their initial photosynthetic product is a 3-carbon molecule, whereas C_4 plants first product is that of a 4-carbon molecule. Ribulose bisphosphate carboxylase-oxygenase (Rubisco) is an enzyme responsible for catalyzing the configuration of two phosphoglycerate molecules when CO_2 is the dominant substrate. This is necessary for photosynthesis to take place and therefore plant growth is dependent on $[CO_2]/[O_2]$ ratios. C4 plants demonstrate a biochemical enhancement of C_3 photosynthesis, in which the Rubisco oxygenase activity is reduced and thus the system is more efficient [5]. A variety of such qualities demonstrated by lignocellulosic plants including; growth rate, ability to grow amidst varying climates, water consumption and cellular composition, in addition to soil and nutrient requirements, have lead us to view certain species as more desirable for domestication and thus modern day food crops. Although, it has not been until recently that these lignocellulosic plants have been considered as a solution for the ever-increasing demand for a renewable fuel source. These same characteristics are being considered when selecting plants, such as switchgrass and maize, for biofuels development [18].

In recent years, the government has taken a more active role in furthering the development of renewable fuel sources. In fact, recent mandates have been implemented requiring the production of 136 billion liters of renewable fuels, 79 billion of which are expected to be cellulosic biofuels, by 2022 in the United States alone [16]. (Figure 2) In order to accomplish such a feat, several efforts have been made to improve the efficiency of the cellulosic biofuels production process.

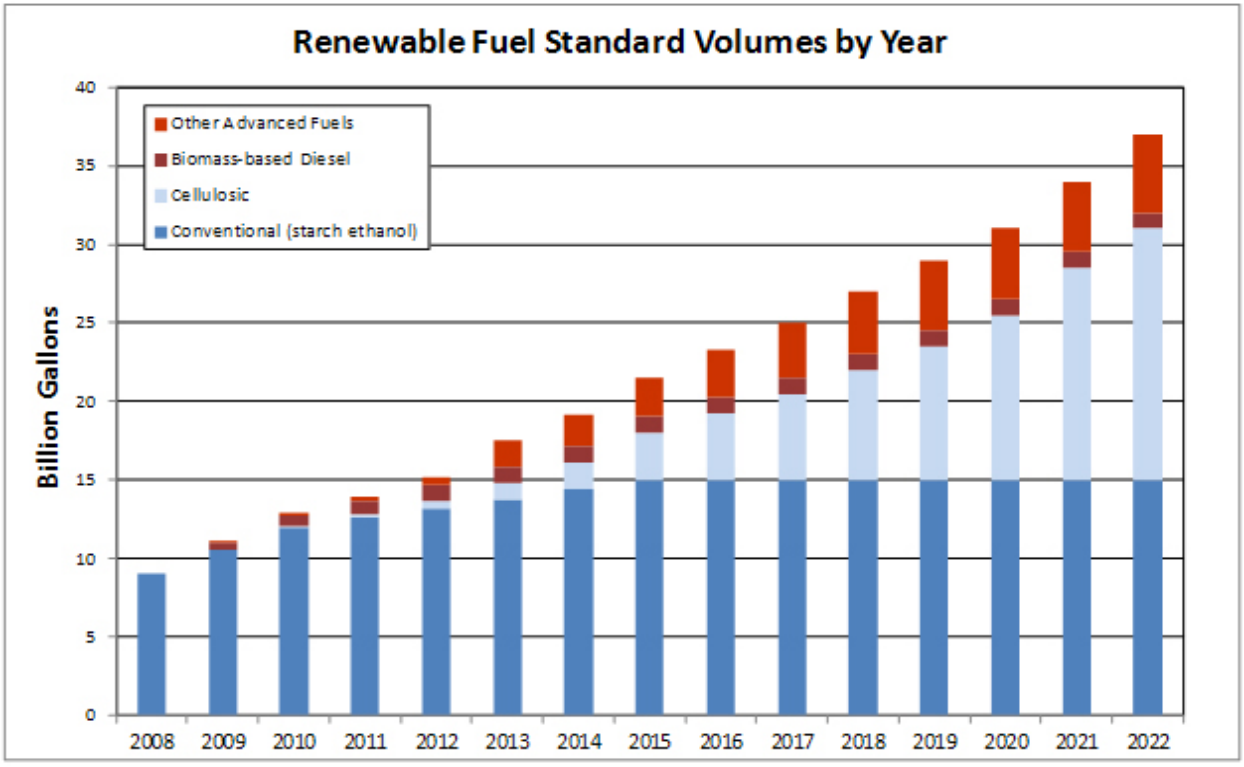

Figure 2. Energy Independence and Security Act of 2007 (EISA) Renewable Fuel Standards. Image shows yearly renewable fuel standard volume requirements for each renewable fuel category. Volume requirements are updated by the Environmental Protection Agency each year based on fuel availability [16].

As previously described, the complexity of the cellulose microfibrils and the surrounding hemicelluloses provide a robustness in C_3 and C_4 plants that has stymied the growth of the cellulosic biofuels industry. Current techniques, for obtaining the sugars entombed within the lignocellulosic plant mass, entail a two-step method in which polysaccharides are hydrolyzed into fermentable sugars by means of acid catalysts. After these sugars have been liberated from the lignocellulosic biomass, they are converted into fuels by means of a variety of methods, many of which involve microscopic organisms and/or costly enzymes. Biochemical advancements are slowly allowing for the replacement of acids with imidazolium-based ionic liquids. It has been shown that such ionic liquids, when enhanced with anions such as chloride, alkyl phosphate and acetate, can facilitate heightened levels of enzymatic hydrolysis of

lignocellulosic biomass [14]. Although these methods may appear as though they would be an attractive alternative to previous inefficient methods, numerous pitfalls remain unchanged. Not only do these methods require multiple phases, but they also produce large quantities of waste fluids and there remains the fact that this method is expensive.

A recently proposed alternative to multi-step processes, is a technology referred to as Consolidated BioProcessing (CBP). CBP involves the exclusion of pricey enzyme additives and merges the two-step process into one low-cost configuration where cellulase enzymes, cellulose degradation and fermentation of sugars takes place simultaneously in a bioreactor [5]. Studies estimate that CBP will allow for a 25% reduction in the cost of biofuel production due to elimination of enzyme expenses [9]. This technology currently involves exploitation of the metabolic capabilities of the microbial organism, *Clostridium phytofermentans.*

Clostridium phytofermentans is an anaerobic mesophile, isolated from the forest soil near the Quabbin Reservoir in Massachusetts, U.S.A. *C. phytofermentans* is a Gram positive member of Cluster XIVa, recently named as the family Lachnospiracea. It demonstrates the ability to metabolize the hexose and pentose components of lignocellulosic biomass, fermenting them to produce primarily ethanol and hydrogen, as well as several other compounds of interest [10]. Not only does *C. phytofermentans* ferment a variety of sugars at an exceptional rate, but it has also been shown to metabolize them simultaneously [6]. Such qualities have prompted the development of *C. phytofermentans* as a biocatalyst for the production of commercial biofuels.

The *C. phytofermentans* genome is complete (Accession number NC_010001) and possesses three loci that encode bacterial microcompartments (BMCs). Of these three loci, one particular locus contains several genes homologous to that of a BMC locus within the *R. inulinivorans* genome. Upon further analysis of the *C. phytofermentans* genome, it has been

shown that this locus influences the configuration of the fucose and rhamnose metabolic pathways including the production of BMCs [10]. (Figure 3) As previously noted, the primary products of *C. phytofermentans* are typically ethanol and hydrogen, however growth of *C. phytofermentans* cultures on media amended with fucose and rhamnose have been shown to produce higher levels of propanol and propionate.

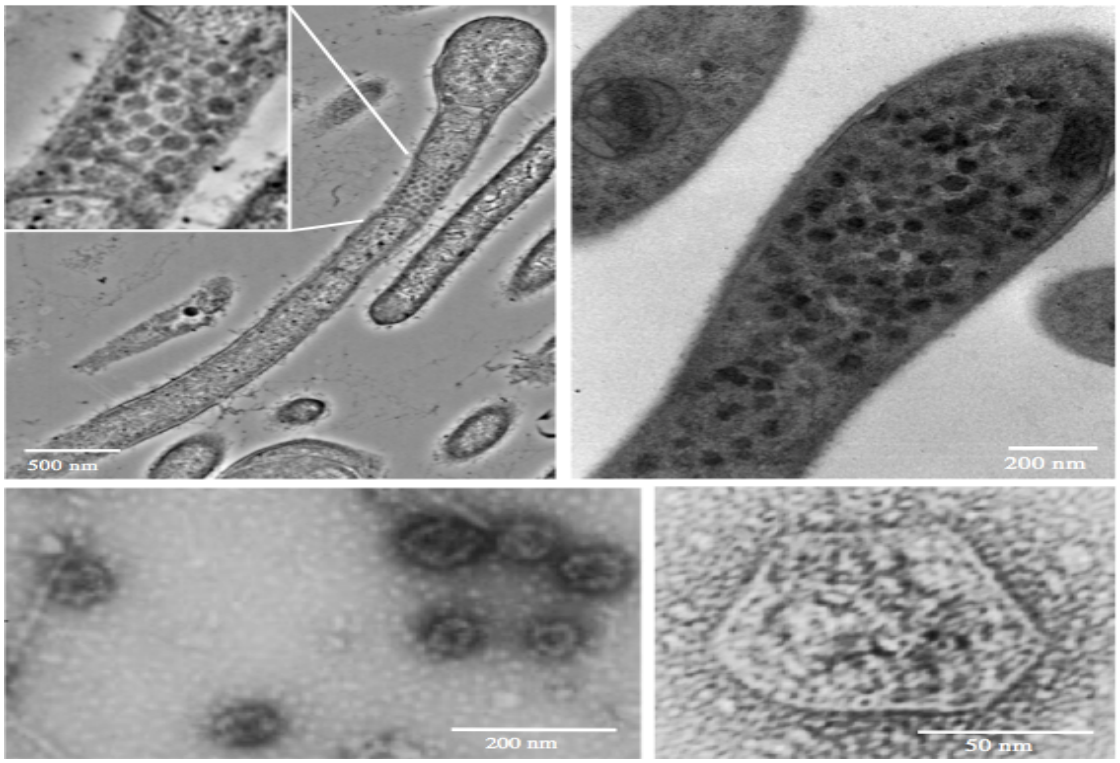

Figure 3. Transmission Electron Micrographs. Images show cross sections of mature *C. phytofermentans* cells in fucose (top left) and rhamnose (top right) displaying several polyhedral structures. (Bottom left) Cells cultured on fucose, showing flagella and BMCs. (Bottom right) BMC from cells cultured on fucose that were sheared of flagella [10].

BMCs are referred to by a variety of terms and possess a variety of components resulting in an assortment of functions. Typically, BMCs range from 80-200nm in size and maintain a polyhedron shape. Current studies assume that these proteinaceous, cyclical hexamers form a

small central pore and are composed of thousands of enzyme and protein subunits, however they lack the nucleic acid and lipid bilayer of viruses and eukaryotic organelles respectively [17]. Although various BMC shells are compiled of analogous proteins due to a conserved protein domain, BMCs do encapsulate a diverse collection of enzymes and thus are responsible for an assortment of metabolic pathways. For example; one form of BMC, carboxysomes, were first observed over 50 years ago in the cytoplasm of cyanobacteria and have since been found in several species of chemoautotrophic bacteria [4]. Carboxysomes carry out the CO_2 fixation process by way of Rubisco, and include carbonic anhydrase [3]. Other BMCs such as the propanediol utilization microcompartment (Pdu) and the ethanolamine utilization microcompartment (Eut) are conditionally produced by *S. enterica* in the presence of 1,2-propanediol, which serves as the carbon source, and are incorporated in its vitamin B_{12}-dependent degradation process [1] [2].

The BMCs produced by *C. phytofermentans* during growth on fucose and rhamnose isolate the metabolism of these substrates presumably to increase metabolic efficiency and protect the cell from lethal metabolic intermediates [13]. The fucose and rhamnose metabolic pathways converge at the production of lactaldehyde from which dihydroxyacetone phosphate (DHAP) and 1,2-propanediol are yielded. It is assumed that DHAP then follows the glycolytic pathway and the resultant products include ethanol, acetate, lactate and formate. Simultaneously, 1,2-propanediol is also further metabolized to generate propanol and propionate as the result of a series of enzymatic reactions. A custom Affymetrix microarray design permitted the measurement of expression values for the genes encoded on this BMC locus. Transcript levels of cultures grown on fucose and rhamnose have been extensively studied due to the highly expressed genes on the BMC locus. D-arabinose portions of this microarray data remained

unanalyzed and have been assessed in the present study as a result of similar expression intensities.

D-arabinose is an uncommon 5-carbon monosaccharide found in biopolymers and has exhibited resistance to digestion by most known microbes [15]. Mapping of the D-arabinose metabolic pathway in *E. coli* initially took place in the 1970's, however little attention has been designated to D-arabinose metabolism in *E. coli*, or any known microbe since the work of LeBlanc and Mortlock in 1971 [7]. Much like fucose, arabinose is more prevalent in nature in its L- form, however the BMC genes of *C. phytofermentans* were not highly expressed in cultures grown on L-arabinose. Therefore, the primary focus of the present study will be on preliminary exploration of D-arabinose metabolism in *C. phytofermentans*. Based on the genomic expression data described above, it was hypothesized that BMCs are produced and play an active role in metabolism of D-arabinose, in *C. phytofermentans.*

CHAPTER II

MATERIALS & METHODS

2.1 Microarray Design and Analysis

The microarray dataset analyzed in the present study was produced by methods noted by Petit, et.al. in, "Involvement of a Bacterial Microcompartment in the Metabolism of Fucose and Rhamnose by Clostridium phytofermentans." (See "Materials and Methods" p. 3) Further analytical proceedings were carried out using a combination of bioinformatics methods, allowing for the comparative visualization of expression values. Software included; Mev: MultiExperiment Viewer, Microsoft Excel and "R".

2.2 Growth Rate Analysis

GS-2 media containing; yeast extract, urea, KH_2PO_4, Na_2HPO_4, trisodium citrate dehydrate, L-cysteine hydrochloride monohydrate, resazurin and milliQ water, was made and adjusted to a pH of 7. The media was then distributed in 10.0mL volumes into 50.0mL tubes and amended with 0.3% glucose, fucose, rhamnose or D-arabinose. The sterilized tubes were flushed with nitrogen gas to create an anaerobic environment and inoculated with 200.0µL *C. phytofermentans* culture, from freezer stock. Cultures were transferred to fresh media on four separate occasions. After the 5^{th} transfer, cultures were monitored and optical densities were recorded with a ThermoSpectronic 3.33183 spectrophotometer at a 660nm wavelength, every 6

hours over a period of 54 hours.

2.3 Protein Analysis & Proteomics

1.0L cultures were grown anaerobically in 2.0L flasks. GS-2 media was amended with 0.3% fucose and cultures were incubated at 32˚C until an optical density of ~0.400 at 660nm was obtained. Cells were harvested by centrifugation at 9,000rpm for 10 minutes at 4˚C. The supernatant was aspirated and cells were resuspended in 40.0mL ice-cold TE buffer. Flagella structures were sheared from the cells by running the solution in a warring blender for 2 minutes. The solution was transferred to a 500.0mL ultra-centrifuge tube and the volume was adjusted to 400.0mL with ice-cold TE buffer. Cells were once again harvested by centrifugation at 9,000rpm for 10 minutes at 4˚C. The supernatant was discarded and pellets were weighed to determine lysis-buffer requirements. Supplemented Cellytic B2X Lysis-Buffer (catalog number MFCD02097909) was used to resuspend the cells and the solution was transferred to a 50.0mL centrifuge tube. The solution was incubated at room temperature for 40 minutes with continuous shaking. Unlysed cells and cellular debris were pelleted via centrifugation at 12,000rpm for 20 minutes at 4˚C. This supernatant was reserved and stored on ice. The pellet was centrifuged again at 12,000rpm for 20 minutes at 4˚C. Supernatants were pooled and centrifuged at 30,000rpm for 90 minutes at 4˚C to pellet the organelles. This pellet, containing the bacterial microcompartments, was resuspended in TEMP buffer and yet again pelleted at 30,000rpm for 90 minutes at 4˚C. The supernatant was discarded and the remaining pellet was resuspended in 10.0mL TEMP buffer. Remaining debris was pelleted out by centrifugation at 12,000rpm for 10 minutes at 4˚C and cells were enriched by concentrating to 1.0mL. This was achieved by

applying the sample to a 100k Vivaspin tube (catalog number 1203039VS) and centrifuging for 10 minutes at 3,220rcf at 4˚C. The 1.0mL sample was divided into 250.0µL aliquots and stored at -80˚C.

Sample concentration was determined with a Fischer Scientific Nanodrop 2000 and 10µg was run in duplicate on BioRad's precast, Mini-Protean TGX polyacrylamide gels. (catalog number 456-9034) A standard Quick Coomassie Blue Staining method was used to determine the presence of protein banding and 100µg of sample were sent to the University of Massachusetts Medical School, Worcester. Mass spectrometry methods were performed and proteomic data was analyzed with Scaffold3 software.

2.4 High-Performance Liquid Chromatography

Triplicate 50.0mL cultures were grown anaerobically in 100.0mL serum bottles on GS-2 growth medium amended with 0.3% substrates including glucose, fucose, rhamnose and D-arabinose. 1.0mL samples were collected over a course of eight time-points, based on their growth rates, and purified with 0.22µm syringe filters prior to freezing at -20˚C in 2.0mL micro-centrifuge tubes. Samples were thawed over ice and 200.0µL aliquots were pipetted into a flat-bottomed 96-well plate with a low evaporation lid. Samples were run for 30 minutes each at 30˚C on a Shimadzu high-performance liquid chromatography system, including an Aminex HPX-87H Ion Exclusion Column, where 5.0mM sulfuric acid served as the running buffer. Samples were then processed using LC Solution Real Time Analysis software. Standards included; glucose, rhamnose, fucose, arabinose, lactate, formate, acetate, propionate, propanol, ethanol and 1,2-propanediol, as well as 3 hydroxypropanoic acid and 1,3-propanediol. Standard curves were calculated and applied to each sample to produce a concentration based on retention

time and area under the corresponding peak.

2.5 Quantitative Polymerase Chain Reaction (qPCR)

10.0mL Cphy cultures were grown in triplicate at 32˚C on GS-2 media adjusted to a 0.3% concentration of either glucose or D-arabinose. Cells were harvested via centrifugation during mid logarithmic growth phase (Optical density: ~0.850 on glucose and ~0.300 on D-arabinose at 660nm). Cells were then pelleted by centrifugation at 4,000rpm for 10 minutes and the supernatant was discarded. An RNA extraction was performed by re-suspending pelleted cells in 2mL N-lauroyl sacrosine solution, centrifuging for 5 minutes at room temperature, at 12,000 rpm. Again the supernatant was aspirated and discarded. The cells were then submersed in 1.0mL of SDS and heated to 100˚C for 5 minutes. This solution was then diluted with 9.0mL Diethyl Polycarbonate-treated water. The samples were pelleted once more via combined centrifugation. The first spin was done at 4,000rpm for fine minutes and pellets were transferred to 2.0mL micro-centrifuge tubes where they were spun an additional 5 minutes at 12,000rpm. The supernatant was discarded and pellets were resuspended in 100µL of 0.001% SDS. A portion of the samples were run on a 1% agarose gel at 120V for 30 minutes to confirm the presence of 16S and 23S subunits. The remaining sample was stored at -80˚C.

Qiagen's QuantiTect Reverse Transcription Kit (catalog numbers 205310, 205311, 205313 and 205314) was used to eliminate the genomic DNA and synthesize complimentary DNA from the consequential RNA template. Treating duplicate samples with all components except the reverse transcriptase produced a negative control for each sample. PCR was then run with 1.0µL of purified cDNA and GAPDH forward and reverse primers. The resulting PCR

products were run on a 1.5% agarose gel for 30 minutes at 120V to confirm that all genomic DNA had been eliminated. qPCR standards were produced by running 1.0µL of purified genomic DNA through a PCR cycle with forward and reverse primers for the four genes chosen for quantitative analysis. These primers were designed by means of Primer3Plus based on sequences obtained from the NCBI database. PCR products were run on a 1.5% agarose gel for 30 minutes at 120V and collected by gel extraction. Extracted samples were purified with Qiagen's QIAquick Gel Extraction Kit (catalog number 28704) and concentrations were obtained by means of a Fischer Scientific Nanodrop 2000. Standards were stored at -80˚C.

qPCR was run on a DNA Engine Opticon 2, with Bio-Rad's iTaq Universal SYBR Green Supermix. (catalog number 172-5121) Each reaction included 7.8µL milliQ water, 10.0µL SYBR Green, 0.6µL of each primer (forward and reverse), and 1µL of cDNA or the corresponding standard. Standard values were converted from nanograms to gene copy number using the following equation: gene copies = (DNA concentration [ng/µL])(1.0g/1,000^{3}ng)(1.0mol bp DNA/330g DNA) X (6.023E023 bp/mol bp)(1 copy/genome or plasmid size [bp]) X (volume of template [µL]) Calculation estimates assume an average molecular weight of 330 per nucleotide in single-stranded DNA [11]. Each set of samples and standards was run twice to serve as a mechanical replicate and samples were run on a 1.5% agarose gel to confirm the absence of errors such as primer-dimers. Data was analyzed via Opticon Monitor 3 software.

CHAPTER III

RESULTS

While analysis of unpublished microarray data showed amplified expression values during growth on fucose and rhamnose, heightened gene expression could also be observed when cultures were grown with D-arabinose serving as the sole carbon source. On average, expression of D-arabinose cultures exceeded that of glucose cultures by greater than 23%. (Figure 4)

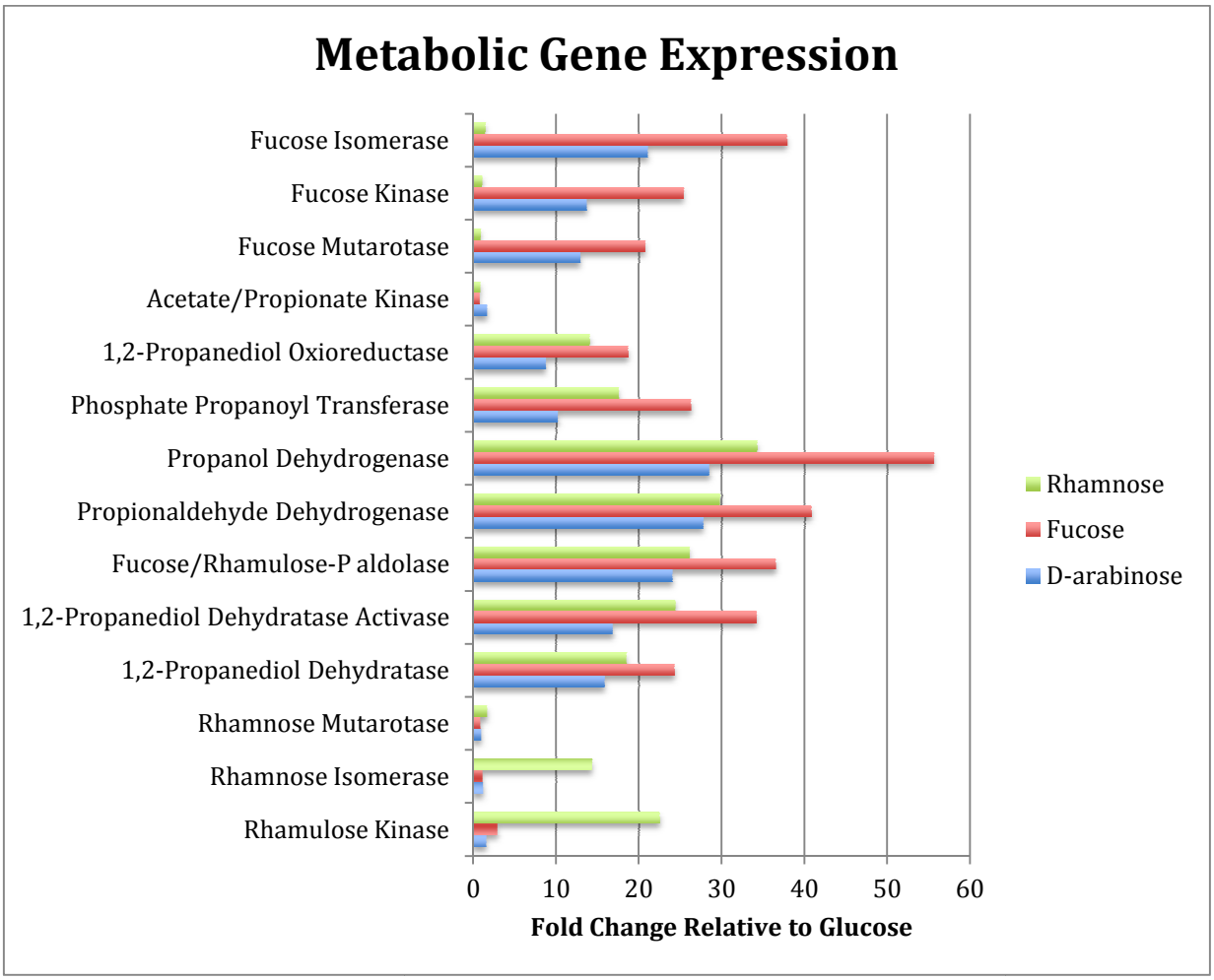

Figure 4. Bacterial Microcompartment Gene Expression. Cultures grown with fucose, rhamnose or D-arabinose as their sole carbon source show elevated expression of BMC genes, relative to glucose [10].

Growth curves were recorded as a preliminary dataset. Based on the infc
obtained, cells were harvested at mid-logarithmic growth phase for all experime
Cultures grown on fucose reached higher optical densities throughout the duratic
growth cycle. However, both fucose and D-arabinose cultures reached their pea
at approximately 48 hours. (Figure 5) Growth of cultures amended with glucose
reached optical densities greater than 1.0 and peaked between 24 and 30 hours. (
data)

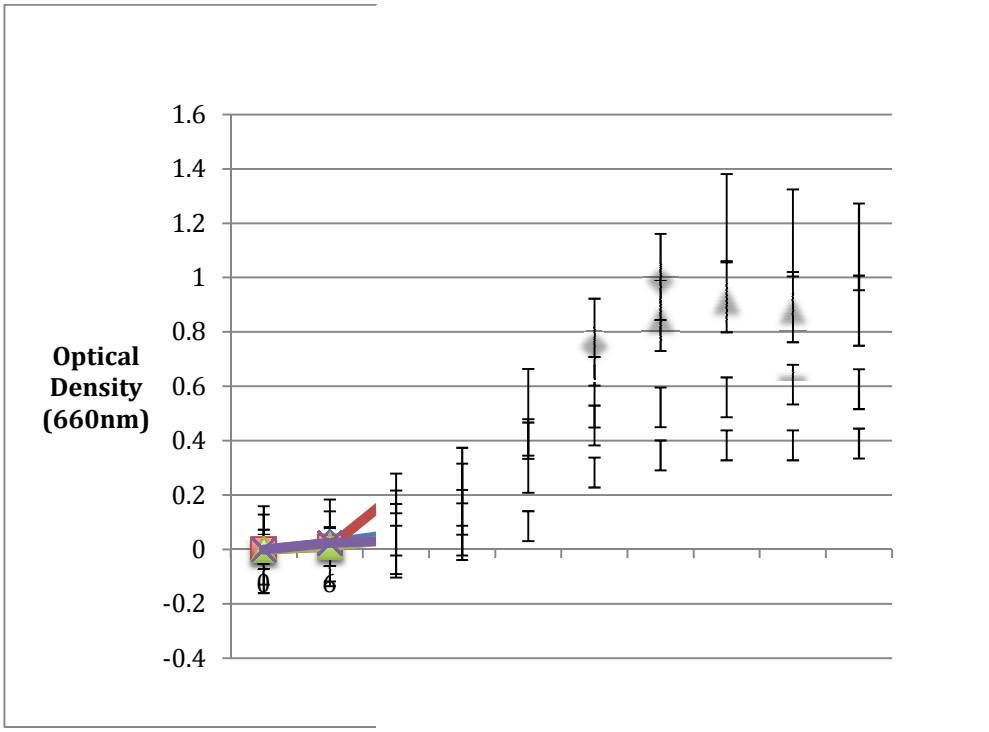

Figure 5. Growth rates of *C. phytofermentans* in media amended with glucos rhamnose or D-arabinose. Optical density was recorded every 6 hours for a tot Cultures were grown in triplicate to serve as a biological replicates. Glucose, rh fucose amended cultures consistently reached higher optical densities than cultu D-arabinose as their sole carbon source.

Based on the significant increase in gene expression on fucose, rhamnose arabinose, high performance liquid chromatography (HPLC) was run to determi

subsequent metabolic products and revealed a variety of sample components. Cultures with fucose or rhamnose serving as their sole carbon source produced a parallel array of compounds, throughout the duration of their growth cycle. Cultures grown on glucose or D-arabinose likewise produced a similar distribution of compounds throughout the extent of their growth cycle. However, multiple compounds produced during growth on fucose and rhamnose were not present when glucose or D-arabinose served as the sole substrate. (Figure 6) Propanol and propionate production commenced approximately halfway through the growth cycle and more than tripled by the time fucose and rhamnose cultures reached stationary phase. All cultures demonstrated a direct correlation between substrate consumption and product yield.

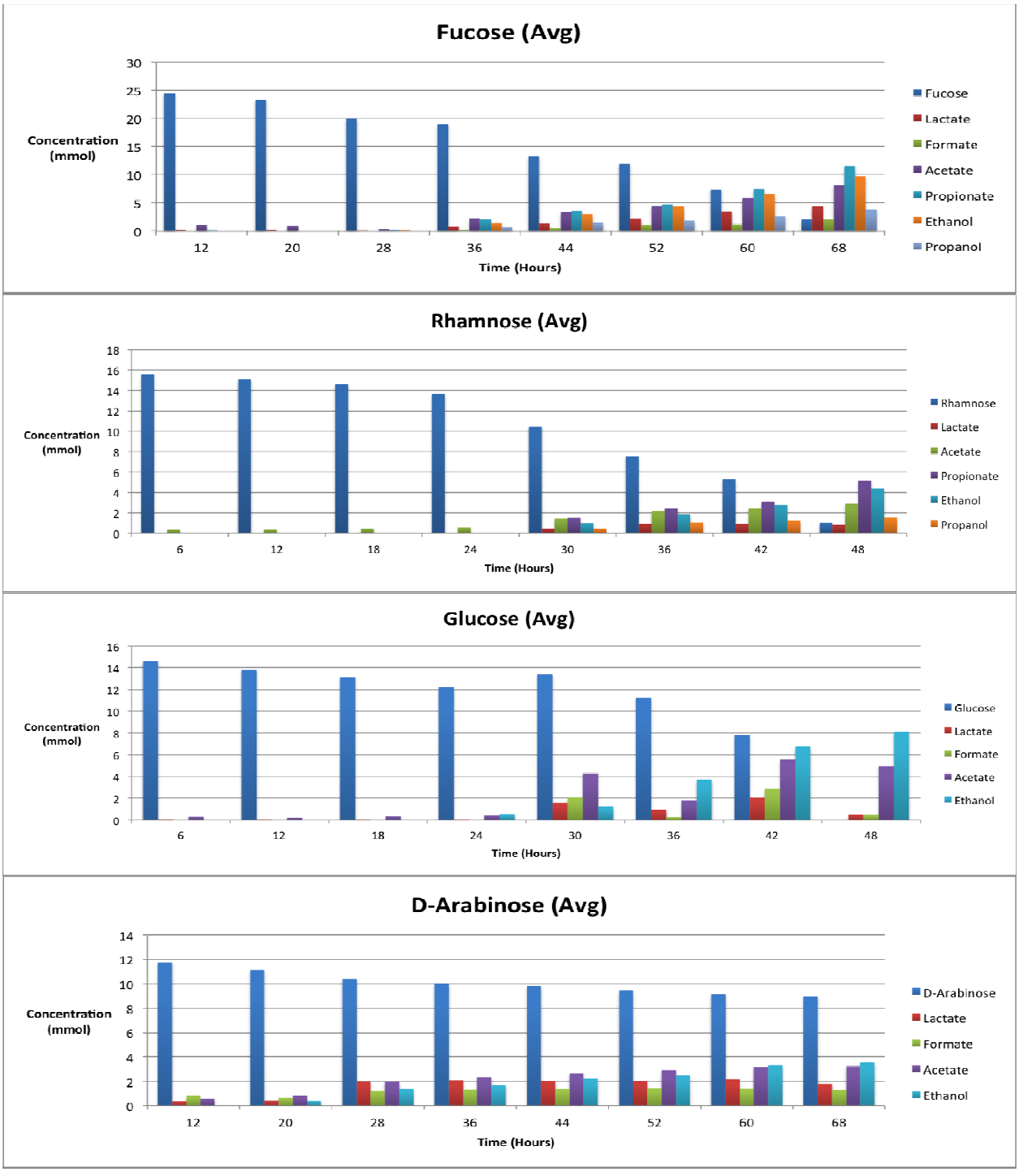

Figure 6. High Performance Liquid Chromatography. 50.0mL Cphy cultures grown on GS-2 media amended with fucose, rhamnose, glucose or D-arabinose, respectively, were assessed for the production of the following: lactate, acetate, formate, ethanol, 1,2-propanediol, 1,3-propanediol, propanol, propionate and 3 hydroxypropionate.

Real-time polymerase chain reaction analysis was completed to evaluate the relative transcript abundance of four genes, shown to be highly expressed on D-arabinose cultures, which are encoded on the BMC locus of interest. These genes included Cphy_3155; L-fucose isomerase, Cphy_1147; L-rhamnose isomerase, Cphy_1174; pyruvate formate-lyase and Cphy_1184; ethanolamine utilization protein EutN/carboxysome structural protein Ccml. Based on previously depicted gene expression values, Cphy_3155, Cphy_1174 and Cphy_1184 from D-arabinose cultures were hypothesized to demonstrate elevated transcript abundance relative to glucose while it was anticipated that Cphy_1147 would show a significantly decreased relative transcript abundance. Contrary to the suggested theory, although Cphy_3155 and Cphy_1147 abundance reflected the microarray data, transcript abundance of Cphy_1174 and Cphy_1184 was greater than 40% less in D-arabinose cultures than in the corresponding glucose cultures. (Figure 7)

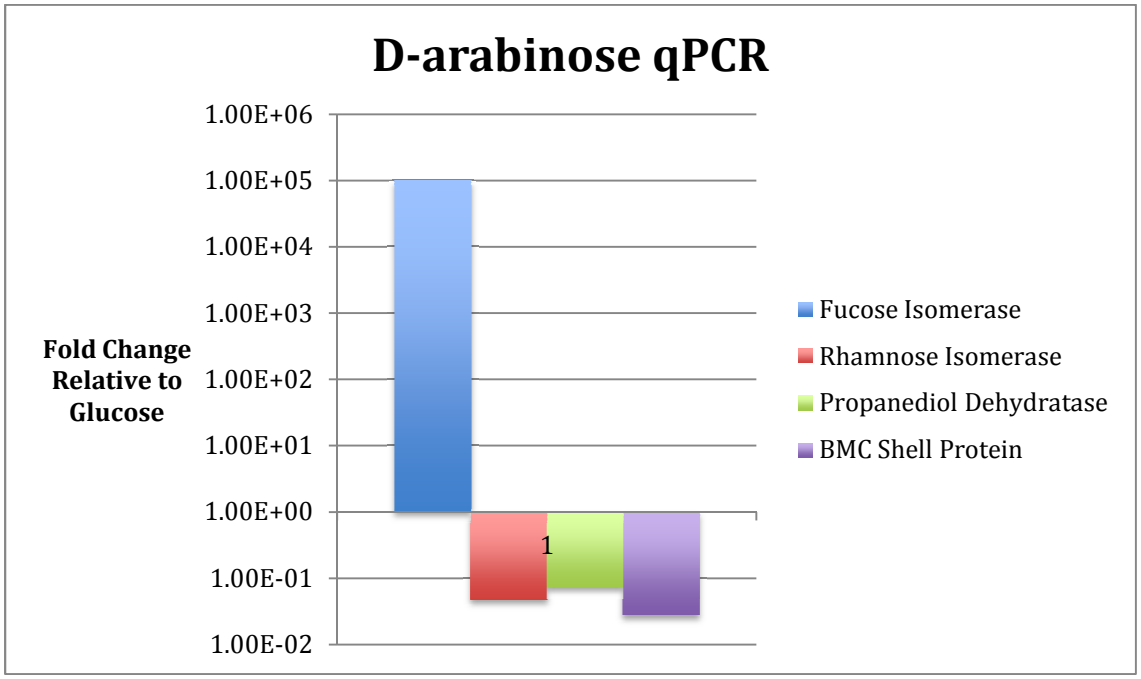

Figure 7. Real-time Polymerase Chain Reaction (qPCR). Real-time polymerase chain reaction analysis of *C. phytofermentans* cultures grown on glucose and D-arabinose allowed for quantitation of relative transcript abundance. The four genes analyzed included: Cphy_3155; L-fucose isomerase, Cphy_1147; L-rhamnose isomerase, Cphy_1174; pyruvate formate-lyase and Cphy_1184; ethanolamine utilization protein EutN/carboxysome structural protein Ccml.

CHAPTER IV

DISCUSSION

C. phytofermentans has exhibited boundless potential in the realm of renewable biofuel production. The ability of *C. phytofermentans* to metabolize the tenacious hexose and pentose components of lignocellulosic biomass, has lead to further examination of the molecular and cellular means by which such tasks are performed in an effort to enhance the efficiency of fermentation, the ulterior motive being that of industrial application. In order to exploit the assets of this unique microbe, a greater level of understanding must be obtained. In an effort to expand upon the current knowledge base, extensive analysis of gene expression data has directed the focus to three BMC loci, one of which is conditionally expressed resulting in the production of BMCs in the presence of fucose, rhamnose and, as is assessed in this study, potentially D-arabinose. A scatterplot of gene expression on D-arabinose and fucose relative to glucose shows a closely entwined transcriptional response of all of the genes that are hypothesized to control the metabolic pathways for both sugars. (Figure 8) As data is collected in regards to these pathways this scatterplot can be used to further assess the transcriptional response of each sugar.

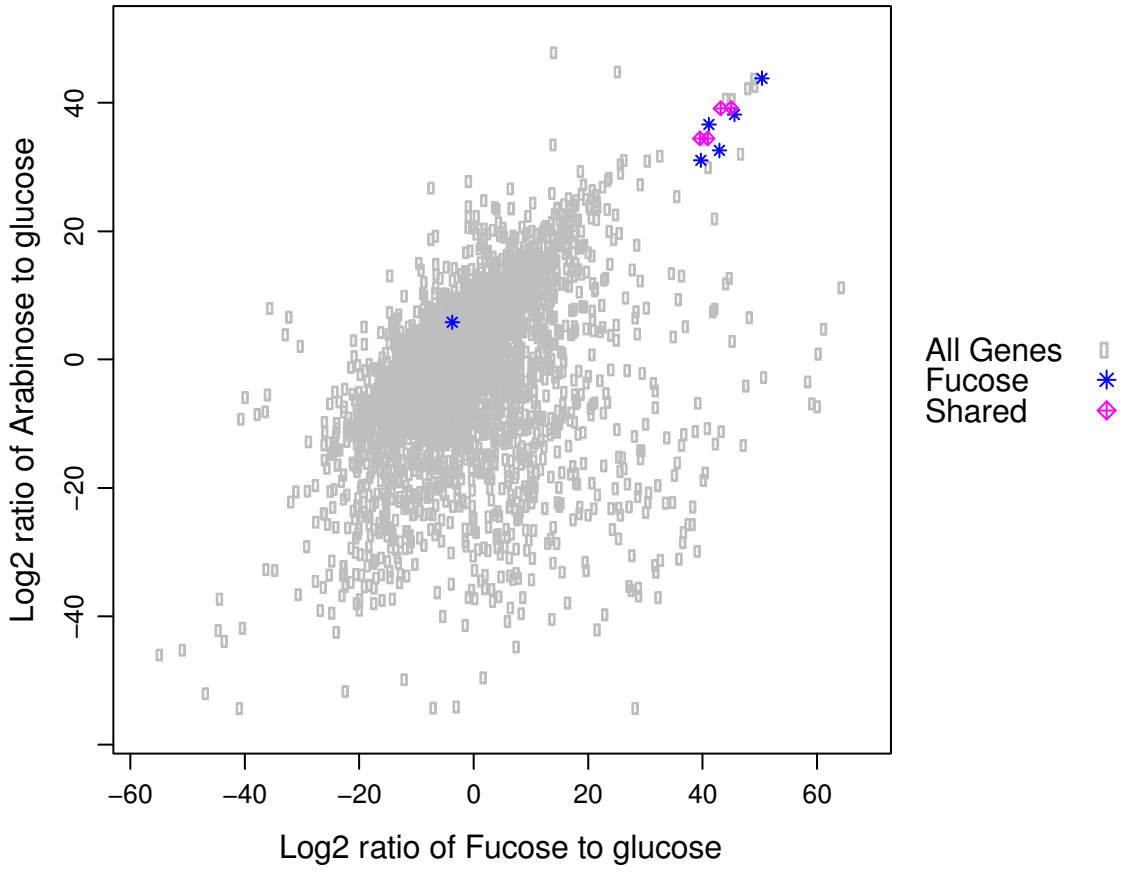

Figure 8. Scatterplot of gene expression values on D-arabinose and fucose. Expression values shown are relative to glucose expression values. D-arabinose values are from the present study and fucose values are provided by Petit et.al. 2013 [10].

Two models have been proposed to map the metabolism of D-arabinose by *C. phytofermentans*. The first of which assumes that BMCs are indeed produced and results in the production of 1,3-propanediol, 3 hydroxypropionate, propanol and propionate. The second pathway, suggests that although the genes are highly expressed and BMCs may be produced, it is the aldolase alone that it required for metabolism and the BMCs therefore go unused resulting in the production of ethanol, lactate, acetate and formate by way of the Embden-Meyerhof pathway. (Figure 9)

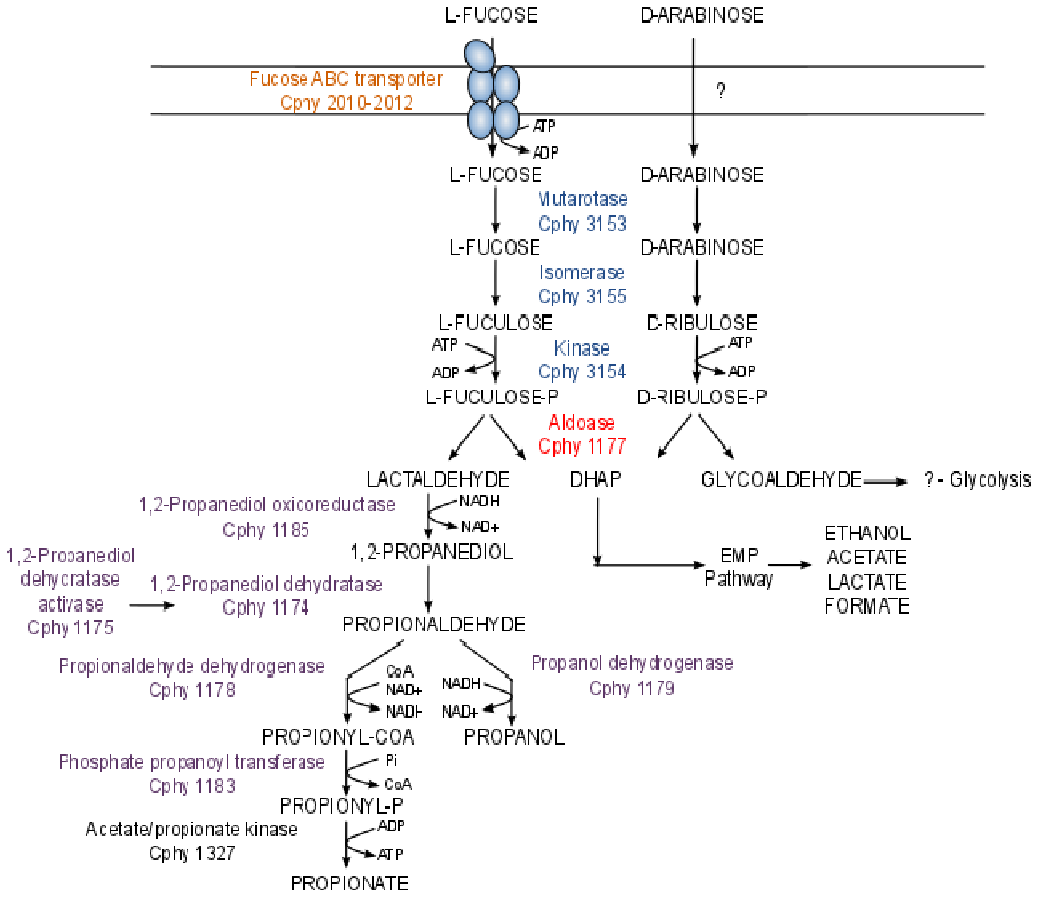

Figure 9. Model of *C. phytofermentans* fucose and D-arabinose metabolic pathways. All genes shown are based on micrarry expression data. The Embden-Meyerhof pathway is designated EMP and "?" marks indicate areas with minimal supporting research. The aldolase hypothesized to play a critical role in the metabolism of both fucose and D-arabinose is indicated in red.

Based on HPLC analysis of growth of *C. phytofermentans* on glucose, fucose, rhamnose and D-arabinose over the course of a full cell cycle, it has been determined that 1,3-propanediol, propanol, 3 hydroxypropionate, propionate and 1,2-propanediol were not produced suggesting that D-arabinose is not metabolized through the BMCs. The production of significant quantities of ethanol, acetate, lactate and formate indicate that D-

arabinose metabolism likely follows the second pathway propositioned and utilizes the EMP pathway.

Real-time polymerase chain reaction was performed on select genes in an effort to reinforce the theories proposed based on the HPLC findings. Low transcript abundance of Cphy_1147, the rhamnose isomerase gene, was anticipated due to the low expression values depicted in the microarray dataset. Cphy_3155, the fucose isomerase gene, also produced transcript levels that reflected gene expression data. This supports the theory that the D-arabinose and fucose metabolic pathways run parallel through the initial phases of digestion. Cphy_1174, the 1,2-propanediol dehydratase gene, unlike the first two genes, demonstrated an indirect correlation between high microarray expression values and low relative transcript abundance. Although it is unclear why the genes are highly expressed if the cells are not producing several copies of the transcript, this may support the proposed model in which 1,2-propanediol dehydratase does not play an active role in the metabolism of D-arabinose. Lastly, Cphy_1184, the carboxysome structural protein Ccml gene, much like Cphy_1174 displayed relative transcript abundance levels inversely reflective of gene expression values. This indicates that BMCs, although genetically encoded for and highly expressed, may in fact not be produced when D-arabinose serves as the solitary carbon source.

CHAPTER V

FUTURE DIRECTIONS

In an effort to devise an explanation for this conundrum further qPCR studies involving the analysis of Cphy_1177, the aldolase gene, have been proposed. Prior to exploring D-arabinose metabolism, attentions were focused on purifying BMCs from *C. phytofermentans* cultures grown on fucose. Although pure preparation of BMCs was never achieved, the presence of BMC proteins was confirmed using polyacrylamide gels and samples were subjected to mass spectrometry. Several of the proteins encoded on the BMC locus were identified. (Table 1) In the future, this method could be applied to D-arabinose cultures in an effort to confirm the production of BMCs. If the appropriate proteins can be accounted for by mass spectrometry, then transmission electron microscopy could be applied to capture images of the proteinaceous organelles.

GENE #:	gi #:	PROPOSED GENE FUNCTION:	MW (kDa):	% TOTAL SPECTRA:
1174	160879323	Propanediol dehydratase	97	0.087
1175	160879324	Propanediol dehydratase activator	29.62	
1176	160879325	BMC shell protein (PF00936)	13	0.0066
1177	160879326	Fuculose and rhamulose phosphate aldolase	29	0.035
1178	160879327	Propionaldehyde dehydrogenase	49.57	
1179	160879328	Propanol dehydrogenase	43	0.11
1180	160879329	BMC shell protein (PF00936)	11	0.0058
1181	160879330	BMC shell protein (PF00936)	11	0.0041
1182	160879331	BMC shell protein (PF00936)	17	0.043
1183	332319813	Phosphate propanoyl transferase	23.06	
1184	160879333	Eut_CcmL shell protein (PF03319)	9	0.017
1185	160879334	Propanediol oxidoreductase	48	0.013
1186	160879335	BMC shell protein (PF00936)	19	0.0058
1187	160879336	Transcriptional Regulator	29	0.0025

Table 1. Analysis of BMC Proteins. Mass spectrometry procedures took place at the University of Massachusetts Medical School, Worcester. Values displayed in red indicate estimated molecular weights based on previously published data, however were not confirmed in our analysis.

BIBLIOGRAPHY

1. Beeby, Morgan, Thomas A. Bobik, and Todd O. Yeates. "Exploiting Genomic Patterns to Discover New Supramolecular Protein Assemblies." *Protein Science* (2008): 69+. Web.

2. Bobik, Thomas A. "Bacterial Microcompartments." *Microbe* (2007): 25-31. Web. 8

3. Cheng, Shouqiang, Yu Liu, Christopher S. Crowley, Todd O. Yeates, and Thomas A. Bobik. "Bacterial Microcompartments: Their Properties and Paradoxes." *NIH Public Access* (2008): 1084-095. Web.

4. Drews, G., and W. Niklowitz. "Cytology of Cyanophycea. II Centroplasm and Granular Inclusions of *Phormidium uncinatum*." *Arch Mikrobiol* (1956): n. pag. *PubMed: 13327992*. Web.

5. Ehleringer, James R., and Thure E. Cerling. "C3 and C4 Photosynthesis." *Encyclopedia of Global Environmental Change* (2002): 186+. Print.

6. *The Gas Is Greener: The Future of Biofuels* (2007) (testimony of Susan B. Leschine). Print.

7. LeBlanc, Donald J., and Robert P. Mortlock. "Metabolism of D-Arabinose: A New Pathway in *Escherichia coli*." *Journal of Bacteriology* (1971): 90+. Print.

8. Leschine, Susan B. "Cellulose Degradation in Anaerobic Environments." *Annual Reviews* (1995): 399+. Print.

9. Lynd, L. R., W. H. Van Zyl, J. E. McBride, and M. Laser. "Consolidated Bioprocessing of Cellulosic Biomass: An Update." *Curr. Opin. Biotechnol.* (2005): 577-83. Web.

10. Petit, Else, Greg LaTouf, Maddalena V. Coppi, Thomas A. Warnick, Devin Currie, Igor Romashko, Supriya Deshpande, Kelly Haas, Jesus G. Alvelo-Maurosa, Colin Wardman, Danny J. Schnell, Susan B. Leschin, and Jeffrey L. Blanchard. "Involvement of a Bacterial Microcompartment in the Metabolism of Fucose and Rhamnose by *Clostridium Phytofermentans*." *PLOS One* E54337 8.1 (2013): n. pag. Web. <www.plosone.org>.

11. Ritalahti, Kristi M., Benjamin K. Amos, Youlboong Sung, Qingzhong Wu, Stephen S. Koenigsberg, and Frank E. Loffler. "Quantitative PCR Targeting 16S RRNA and Reductive Dehalogenase Genes Simultaneously Monitors Multiple *Dehalococcoides* Strains." *Applied and Environmental Microbiology* (2006): 2765-774. Print.

12. Rubin, Edward M. "Genomics of Cellulosic Biofuels." *Nature* (2008): n. pag. Print.

13. Sampson, E., and T. Bobik. "Microcompartments for B12-dependent 1,2-propanediol Degradation Provide Protection from DNA and Cellular Damage by a Reactive Metabolic Intermediate." *Journal of Bacteriology* (2008): 2966-971. Web.

14. Sun, Ning, Hanbin Liu, Noppadon Sathitsuksanoh, Vitalie Stavila, Manali Sawant, Anaise Bonito, Kim Tran, Anthe George, Kenneth L. Sale, Seema Singh, Blake A. Simmons, and Bradley A. Holmes. "Production and Extraction of Sugars from Switchgrass Hydrolyzed in Ionic Liquids." *Biotechnology for Biofuels* (2013): n. pag. Web. <www.biotechnologyforbiofuels.com/content/6/1/39>.

15. Tian, J., H. Che, D. Ha, Y. Wei, and S. Zheng. "Characterization and Anti-allergenic Effect of a Polysaccharide from Flower Buds of *Lonicera japonica*." *EPub* (2012): n. pag. Print.

16. USA, (The United States of America). "The Energy Independence and Security Act of 2007." (2007): n. pag. Web.

17. Yeates, Todd O., Christopher S. Crowley, and Shiho Tanaka. "Bacterial Microcompartment Organelles: Protein Shell Structure and Evolution." *NIH Public Access* (2010): 185-205. Web.

18. Zhuang, Qianlai, Zhangcai Qin, and Min Chen. "Biofuel, Land and Water: Maize, Switchgrass or *Miscanthus*?" *Environmental Research Letters* (2013): n. pag. Print.